

1/5/1 (Item 1 from file: 351)  
DIALOG(R) File 351:Derwent WPI  
(c) 2006 The Thomson Corporation. All rts. reserv.

0013314332 - Drawing available  
WPI ACC NO: 2003-401486/200338  
XRPX Acc No: N2003-320179

**Inter-processor communication system for parallel computer system,  
transmits packet requesting translation of write address before  
transmission of data, from source processor to destination processor**  
Patent Assignee: NEC CORP (NIDE)  
Inventor: KANO T; KANO H Y

**Patent Family** (3 patents, 2 countries)

Patent Number	Kind	Date	Application Number	Kind	Date	Update
US 20030005071	A1	20030102	US 2002161636	A	20020605	200338 B
JP 2002366427	A	20021220	JP 2001170377	A	20010606	200340 E
JP 3800037	B2	20060719	JP 2001170377	A	20010606	200648 E

Priority Applications (no., kind, date): JP 2001170377 A 20010606

#### Patent Details

Number	Kind	Lan	Pg	Dwg	Filing Notes
US 20030005071	A1	EN	13	5	
JP 2002366427	A	JA	11		
JP 3800037	B2	JA	13		Previously issued patent JP 2002366427

#### Alerting Abstract US A1

NOVELTY - A transmitter (4) transmits a packet requesting translation of a write address before transmission of data, from a source processor to a destination processor. An address translating circuit (54) translates the write address to a physical address. A data write circuit (57) uses the result of the address translation as the write address for writing the data transmitted subsequent to the packet.

DESCRIPTION - An INDEPENDENT CLAIM is also included for an inter-processor communication method.

USE - For communication between processors connected through network, in parallel computer system.

ADVANTAGE - Eliminates need for address translation at the initial normal packet with data arrival and hence reduces time required for address translation in the receiver. Also reduces capacity of the memory for holding the address translation result and reduces overhead of the address translation in the receiver.

DESCRIPTION OF DRAWINGS - The figure shows a block diagram illustrating the inter-processor communication system.

- 1-1 - 1-n processors
- 4 transmitter
- 5 receiver
- 53 write address register
- 54 address translating circuit
- 55 translation table
- 57 data write circuit
- 100 network

**Title Terms/Index Terms/Additional Words:** INTER; PROCESSOR; COMMUNICATE; SYSTEM; PARALLEL; COMPUTER; TRANSMIT; PACKET; REQUEST; TRANSLATION; WRITING; ADDRESS; TRANSMISSION; DATA; SOURCE; DESTINATION

#### Class Codes

International Classification (Main): G06F-012/02  
(Additional/Secondary): G06F-012/06, G06F-015/163, G06F-015/177  
International Classification (+ Attributes)  
IPC + Level Value Position Status Version

G06F-0012/02	A	I	L	B	20060101
G06F-0012/06	A	I	L	B	20060101
G06F-0012/08	A	I	L	B	20060101
G06F-0012/10	A	I	F	B	20060101

G06F-0015/17 A I R 20060101  
G06F-0015/17 A I L B 20060101  
G06F-0015/16 C I R 20060101  
G06F-0015/16 C I L B 20060101  
US Classification, Issued: 709213000

File Segment: EPI;

DWPI Class: T01

Manual Codes (EPI/S-X): T01-H07C7; T01-M02A; T01-M02C

(19)日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11)特許出願公開番号

特開2002-366427

(P2002-366427A)

(43)公開日 平成14年12月20日(2002.12.20)

(51)Int.Cl. <sup>7</sup>	識別記号	F I	ターミナル(参考)
G 0 6 F 12/02	5 7 0	G 0 6 F 12/02	5 7 0 K 5 B 0 4 5
12/06	5 3 0	12/06	5 3 0 D 5 B 0 6 0
15/163	6 5 0	15/163	6 5 0 A
15/177	6 7 6	15/177	6 7 6 A

審査請求 未請求 請求項の数18 O L (全 11 頁)

(21)出願番号 特願2001-170377(P2001-170377)

(22)出願日 平成13年6月6日(2001.6.6)

(71)出願人 000004237

日本電気株式会社

東京都港区芝五丁目7番1号

(72)発明者 加納 健

東京都港区芝五丁目7番1号 日本電気株式会社内

(74)代理人 100088812

弁理士 ▲柳▼川 信

Fターム(参考) 5B045 BB02 BB47 BB48 DD07

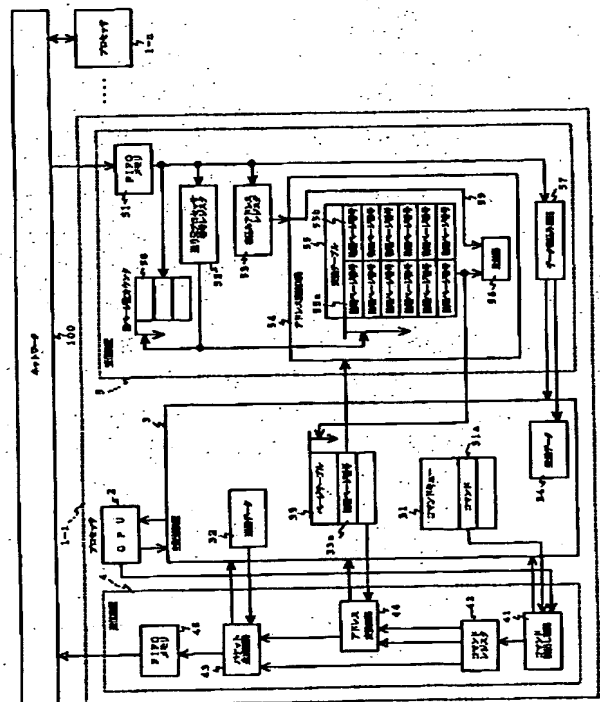
5B060 AB26 AC19 KA01 KA10

(54)【発明の名称】 プロセッサ間通信システム及びそれに用いるプロセッサ間通信方法

(57)【要約】

【課題】 受信装置でのアドレス変換のオーバーヘッドの削減や隠蔽を図ることが可能なプロセッサ間通信システムを提供する。

【解決手段】 送信装置4はコマンドの最初のパケットを送る前に書込みアドレス(アドレス変換パケット)を送る。受信装置5はアドレス変換パケットを受取ると、書込みアドレスのアドレス変換を先行して行う。また、受信装置5はプロセッサ間通信中に発生するアドレス変換を先行して行うことで、ページ毎に発生する送り元及び宛先でのアドレス変換のオーバーヘッドを削減する。さらに、送信装置4は総ページ数を宛先に教えることで、最後のページ以降の無駄なアドレス変換を抑止する。



## 【特許請求の範囲】

【請求項 1】 複数のプロセッサ各々がネットワークによって結合される並列コンピュータのプロセッサ間通信システムであって、宛先プロセッサへのデータ送信時に論理アドレスで指定された前記宛先プロセッサでの書込みアドレスのアドレス変換を依頼するためのパケットを前記宛先プロセッサに送信する送信手段と、送り元プロセッサから送られてきた書込みアドレスを物理アドレスに変換する手段と、その変換結果を前記アドレス変換を依頼するためのパケットの後に送られてくるデータの書込みアドレスとする手段とを前記複数のプロセッサ各々に有することを特徴とするプロセッサ間通信システム。

【請求項 2】 前記送信手段は、前記論理アドレスで指定された送信データの読出しアドレスを物理アドレスに変換する前に前記宛先プロセッサに対して前記書込みアドレスのアドレス変換を依頼するパケットを送信するようにしたことを特徴とする請求項 1 記載のプロセッサ間通信システム。

【請求項 3】 前記送信データを送るパケットにおいて前記書込みアドレスを指定しておき、先に送られたアドレス変換を依頼するパケットに付加された書込みアドレスと前記送信データを送るパケットに指定された書込みアドレスとが異なる時に前記送信データを送るパケットに指定された書込みアドレスを再変換して用いるようにしたことを特徴とする請求項 2 記載のプロセッサ間通信システム。

【請求項 4】 複数のプロセッサ各々がネットワークによって結合される並列コンピュータのプロセッサ間通信システムであって、宛先プロセッサでのパケット受信時に論理アドレスで指定された前記宛先プロセッサでの書込みアドレスを物理アドレスに変換する手段と、その変換結果を記憶する手段と、記憶している変換結果が使えなくなる前に次に必要となるアドレス変換を先行して行う手段とを前記複数のプロセッサ各々に有することを特徴とするプロセッサ間通信システム。

【請求項 5】 複数のプロセッサ各々がネットワークによって結合される並列コンピュータのプロセッサ間通信システムであって、宛先プロセッサへのデータ送信時に論理アドレスで指定された送信データの読出しアドレスを物理アドレスに変換する手段と、その変換結果を記憶する手段と、記憶している変換結果が使えなくなる前に次に必要となるアドレス変換を先行して行う手段とを前記複数のプロセッサ各々に有することを特徴とするプロセッサ間通信システム。

【請求項 6】 複数のプロセッサ各々がネットワークによって結合される並列コンピュータのプロセッサ間通信システムであって、宛先プロセッサでのパケット受信時に論理アドレスで指定された前記宛先プロセッサでの書込みアドレスを物理アドレスに変換する手段と、その変換結果を記憶する手段と、記憶している変換結果が使え

なくなる前に次に必要となるアドレス変換を先行して行う手段と、宛先プロセッサへのデータ送信時に論理アドレスで指定された送信データの読出しアドレスを物理アドレスに変換する手段と、その変換結果を記憶する手段と、記憶している変換結果が使えなくなる前に次に必要となるアドレス変換を先行して行う手段とを前記複数のプロセッサ各々に有し、前記送信データを送るパケットにおいて前記書込みアドレスを指定しておき、先に送られたアドレス変換を依頼するパケットに付加された書込みアドレスと前記送信データを送るパケットに指定された書込みアドレスとが異なる時に前記送信データを送るパケットに指定された書込みアドレスを再変換して用いるようにしたことを特徴とするプロセッサ間通信システム。

【請求項 7】 1つの通信コマンドで複数のパケットによってデータが送られる場合、前記アドレス変換を依頼するパケットをそのコマンドの最初のデータを送るパケットの前だけに送り、その後に前記データを送るパケットだけを送るようにしたことを特徴とする請求項 6 記載のプロセッサ間通信システム。

【請求項 8】 前記変換結果を記憶する記憶手段は、送り元プロセッサ毎に変換結果を管理し、同じ送り元プロセッサに対して現在使用中の変換結果と先行してアドレス変換した結果との 2つの変換結果を記憶するようにしたことを特徴とする請求項 7 記載のプロセッサ間通信システム。

【請求項 9】 前記アドレス変換を要求するパケットにそのコマンドで送る総データ長の受信側でのページ数を指定して送り、前記宛先プロセッサにおいてそのページ数をカウントダウンすることで前記コマンドで書かれる最後のページを予測し、前記最後のページ以降の無駄な先行するアドレス変換を抑止するようにしたことを特徴とする請求項 8 記載のプロセッサ間通信システム。

【請求項 10】 複数のプロセッサ各々がネットワークによって結合される並列コンピュータのプロセッサ間通信方法であって、宛先プロセッサでのパケット受信時に論理アドレスで指定された前記宛先プロセッサでの書込みアドレスのアドレス変換を依頼するためのパケットを前記宛先プロセッサに送信するステップと、送り元プロセッサから送られてきた書込みアドレスを物理アドレスに変換するステップと、その変換結果を前記アドレス変換を依頼するためのパケットの後に送られてくるデータの書込みアドレスとするステップとを前記複数のプロセッサ各々に有することを特徴とするプロセッサ間通信方法。

【請求項 11】 前記書込みアドレスのアドレス変換を依頼するためのパケットを送信するステップは、前記論理アドレスで指定された送信データの読出しアドレスを物理アドレスに変換する前に前記宛先プロセッサに対して前記書込みアドレスのアドレス変換を依頼するパケッ

トを送信するようにしたことを特徴とする請求項10記載のプロセッサ間通信方法。

【請求項12】 前記送信データを送るパケットにおいて前記書込みアドレスを指定しておき、先に送られたアドレス変換を依頼するパケットに付加された書込みアドレスと前記送信データを送るパケットに指定された書込みアドレスとが異なる時に前記送信データを送るパケットに指定された書込みアドレスを再変換して用いるようにしたことを特徴とする請求項11記載のプロセッサ間通信方法。

【請求項13】 複数のプロセッサ各々がネットワークによって結合される並列コンピュータのプロセッサ間通信方法であって、宛先プロセッサでのパケット受信時に論理アドレスで指定された前記宛先プロセッサでの書込みアドレスを物理アドレスに変換するステップと、その変換結果を記憶するステップと、記憶している変換結果が使えなくなる前に次に必要となるアドレス変換を先行して行うステップとを前記複数のプロセッサ各々に有することを特徴とするプロセッサ間通信方法。

【請求項14】 複数のプロセッサ各々がネットワークによって結合される並列コンピュータのプロセッサ間通信方法であって、宛先プロセッサへのデータ送信時に論理アドレスで指定された送信データの読出しアドレスを物理アドレスに変換するステップと、その変換結果を記憶するステップと、記憶している変換結果が使えなくなる前に次に必要となるアドレス変換を先行して行うステップとを前記複数のプロセッサ各々に有することを特徴とするプロセッサ間通信方法。

【請求項15】 複数のプロセッサ各々がネットワークによって結合される並列コンピュータのプロセッサ間通信システムであって、宛先プロセッサでのパケット受信時に論理アドレスで指定された前記宛先プロセッサでの書込みアドレスを物理アドレスに変換するステップと、その変換結果を記憶するステップと、記憶している変換結果が使えなくなる前に次に必要となるアドレス変換を先行して行うステップと、宛先プロセッサへのデータ送信時に論理アドレスで指定された送信データの読出しアドレスを物理アドレスに変換するステップと、その変換結果を記憶するステップと、記憶している変換結果が使えなくなる前に次に必要となるアドレス変換を先行して行うステップとを前記複数のプロセッサ各々に有し、前記送信データを送るパケットにおいて前記書込みアドレスを指定しておき、先に送られたアドレス変換を依頼するパケットに付加された書込みアドレスと前記送信データを送るパケットに指定された書込みアドレスとが異なる時に前記送信データを送るパケットに指定された書込みアドレスを再変換して用いるようにしたことを特徴とするプロセッサ間通信方法。

【請求項16】 1つの通信コマンドで複数のパケットによってデータが送られる場合、前記アドレス変換を依

頼するパケットをそのコマンドの最初のデータを送るパケットの前だけに送り、その後に前記データを送るパケットだけを送るようにしたことを特徴とする請求項15記載のプロセッサ間通信方法。

【請求項17】 前記変換結果を記憶するステップは、送り元プロセッサ毎に変換結果を管理し、同じ送り元プロセッサに対して現在使用中の変換結果と先行してアドレス変換した結果との2つの変換結果を記憶するようにしたことを特徴とする請求項16記載のプロセッサ間通信方法。

【請求項18】 前記アドレス変換を要求するパケットにそのコマンドで送る総データ長の受信側でのページ数を指定して送り、前記宛先プロセッサにおいてそのページ数をカウントダウンすることで前記コマンドで書かれる最後のページを予測し、前記最後のページ以降の無駄な先行するアドレス変換を抑止するようにしたことを特徴とする請求項17記載のプロセッサ間通信方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明はプロセッサ間通信システム及びそれに用いるプロセッサ間通信方法に関し、特に複数のプロセッサがネットワークによって結合された並列コンピュータにおいてデータの読出しアドレスと書込みアドレスとに論理アドレスを使ったプロセッサ間通信方法に関する。

【0002】

【従来技術】従来、プロセッサ間の通信においては、宛先プロセッサでのデータ書込みアドレスを間違えると、宛先プロセッサをシステムダウンさせる可能性がある。そのため、さまざまな手段によって宛先プロセッサを保護している。

【0003】例えば、受信バッファ領域を決めて、プロセッサ間通信ではこの領域以外には書込むことができない方法がある（以下、第1の従来技術とする）。この場合、書込みアドレスは宛先プロセッサの受信装置が決定した物理アドレスか、送り元プロセッサが指定してきた物理アドレスを、宛先プロセッサの受信装置がマスク等によって限定している。

【0004】また、書込みアドレスを物理アドレスではなく、論理アドレスで送り元が指定する方法もある（以下、第2の従来技術とする）。この場合、宛先プロセッサの受信装置がアドレス変換を行って物理アドレスに変換している。この方法では受信装置でアドレス変換して得られる物理アドレスを特定のアドレスの範囲に制限することができるため、宛先プロセッサを保護することができる。仮に、指定された書込みアドレスが不正な場合にはアドレス変換が失敗することになる。

【0005】上述した第1の従来技術では、受信バッファ領域からユーザ領域へのデータのコピーが必要になり、プロセッサ間通信のスループット性能とレイテンシ

10

20

30

40

50

性能とを確保することが難しい。そのため、第2の従来技術の書込みアドレスを論理アドレスで指定する方法が使われることが多い。

【0006】しかしながら、第2の従来技術では指定された論理アドレスから物理アドレスに変換する手段が必要である。このアドレス変換はページを単位に行われ、通常ページサイズは4Kバイト等である。そのため、プロセッサ間通信で大量にデータを送るには、変換テーブルが非常に大きなものになり、受信装置内に全てを保持することができなくなる。したがって、変換テーブル全体を主記憶装置上に実装し、必要に応じて受信装置が主記憶装置から変換テーブルの一部を読み出してアドレス変換を行うことになる。

【0007】そのため、アドレス変換では主記憶装置へのアクセスが発生し、プロセッサ間通信のオーバーヘッドとなる。特に、パケットを受信した時にアドレス変換のために主記憶装置をアクセスすると、パケット内のデータの書込みをアドレス変換後でないと行うことができないため、その間、ネットワークを止めてしまうこととなり、他のプロセッサ間通信にも影響を与える。このようなオーバーヘッドを削減することが、プロセッサ間通信の性能を向上するためには必要となる。

【0008】

【発明が解決しようとする課題】上述した従来のプロセッサ間の通信では、並列計算機システムにおいて宛先プロセッサでの書込みアドレスを論理アドレスで指定した場合、受信装置でのアドレス変換のオーバーヘッドを削減したり、そのオーバーヘッドを隠蔽したりすることが困難である。

【0009】そこで、本発明の目的は上記の問題点を解消し、受信装置でのアドレス変換のオーバーヘッドの削減や隠蔽を図ることができるプロセッサ間通信システム及びそれに用いるプロセッサ間通信方法を提供することにある。

【0010】

【課題を解決するための手段】本発明による第1のプロセッサ間通信システムは、複数のプロセッサ各々がネットワークによって結合される並列コンピュータのプロセッサ間通信システムであって、宛先プロセッサへのデータ送信時に論理アドレスで指定された前記宛先プロセッサでの書込みアドレスのアドレス変換を依頼するためのパケットを前記宛先プロセッサに送信する送信手段と、送り元プロセッサから送られてきた書込みアドレスを物理アドレスに変換する手段と、その変換結果を前記アドレス変換を依頼するためのパケットの後に送られてくるデータの書込みアドレスとする手段とを前記複数のプロセッサ各々に備えている。

【0011】また、本発明による第1のプロセッサ間通信システムの送信手段は、前記論理アドレスで指定された送信データの読み出しアドレスを物理アドレスに変換す

る前に前記宛先プロセッサに対して前記書込みアドレスのアドレス変換を依頼するパケットを送信している。

【0012】さらに、本発明による第1のプロセッサ間通信システムの送信データを送るパケットにおいて前記書込みアドレスを指定しておき、先に送られたアドレス変換を依頼するパケットに付加された書込みアドレスと前記送信データを送るパケットに指定された書込みアドレスとが異なる時に前記送信データを送るパケットに指定された書込みアドレスを再変換して用いる。

10 【0013】本発明による第2のプロセッサ間通信システムは、複数のプロセッサ各々がネットワークによって結合される並列コンピュータのプロセッサ間通信システムであって、宛先プロセッサでのパケット受信時に論理アドレスで指定された前記宛先プロセッサでの書込みアドレスを物理アドレスに変換する手段と、その変換結果を記憶する手段と、記憶している変換結果が使えなくなる前に次に必要となるアドレス変換を先行して行う手段とを前記複数のプロセッサ各々に備えている。

20 【0014】本発明による第3のプロセッサ間通信システムは、複数のプロセッサ各々がネットワークによって結合される並列コンピュータのプロセッサ間通信システムであって、宛先プロセッサへのデータ送信時に論理アドレスで指定された送信データの読み出しアドレスを物理アドレスに変換する手段と、その変換結果を記憶する手段と、記憶している変換結果が使えなくなる前に次に必要となるアドレス変換を先行して行う手段とを前記複数のプロセッサ各々に備えている。

30 【0015】本発明による第4のプロセッサ間通信システムは、複数のプロセッサ各々がネットワークによって結合される並列コンピュータのプロセッサ間通信システムであって、宛先プロセッサでのパケット受信時に論理アドレスで指定された前記宛先プロセッサでの書込みアドレスを物理アドレスに変換する手段と、その変換結果を記憶する手段と、記憶している変換結果が使えなくなる前に次に必要となるアドレス変換を先行して行う手段と、宛先プロセッサへのデータ送信時に論理アドレスで指定された送信データの読み出しアドレスを物理アドレスに変換する手段と、その変換結果を記憶する手段と、記憶している変換結果が使えなくなる前に次に必要となるアドレス変換を先行して行う手段とを前記複数のプロセッサ各々に備え、前記送信データを送るパケットにおいて前記書込みアドレスを指定しておき、先に送られたアドレス変換を依頼するパケットに付加された書込みアドレスと前記送信データを送るパケットに指定された書込みアドレスとが異なる時に前記送信データを送るパケットに指定された書込みアドレスを再変換して用いるようにしている。

50 【0016】また、本発明による第4のプロセッサ間通信システムでは、1つの通信コマンドで複数のパケットによってデータが送られる場合、アドレス変換を依頼す

るパケットをそのコマンドの最初のデータを送るパケットの前だけに送り、その後にデータを送るパケットだけを送るようにしている。

【0017】さらに、本発明による第4のプロセッサ間通信システムのアドレス変換結果を記憶する記憶手段は、送り元プロセッサ毎に変換結果を管理し、同じ送り元プロセッサに対して現在使用中の変換結果と先行してアドレス変換した結果との2つの変換結果を記憶するようにしている。

【0018】さらにまた、第4のプロセッサ間通信システムのアドレス変換を要求するパケットにそのコマンドで送る総データ長の受信側でのページ数を指定して送り、前記宛先プロセッサにおいてそのページ数をカウントダウンすることで前記コマンドで書かれる最後のページを予測し、前記最後のページ以降の無駄な先行するアドレス変換を抑止するようにしている。

【0019】本発明による第1のプロセッサ間通信方法は、複数のプロセッサ各々がネットワークによって結合される並列コンピュータのプロセッサ間通信方法であって、宛先プロセッサでのパケット受信時に論理アドレスで指定された前記宛先プロセッサでの書込みアドレスのアドレス変換を依頼するためのパケットを前記宛先プロセッサに送信するステップと、送り元プロセッサから送られてきた書込みアドレスを物理アドレスに変換するステップと、その変換結果を前記アドレス変換を依頼するためのパケットの後に送られてくるデータの書込みアドレスとするステップとを前記複数のプロセッサ各々に備えている。

【0020】また、本発明による第1のプロセッサ間通信方法の書込みアドレスのアドレス変換を依頼するためのパケットを送信するステップが、前記論理アドレスで指定された送信データの読出しアドレスを物理アドレスに変換する前に前記宛先プロセッサに対して前記書込みアドレスのアドレス変換を依頼するパケットを送信している。

【0021】さらに、本発明による第1のプロセッサ間通信方法の送信データを送るパケットにおいて前記書込みアドレスを指定しておき、先に送られたアドレス変換を依頼するパケットに付加された書込みアドレスと前記送信データを送るパケットに指定された書込みアドレスとが異なる時に前記送信データを送るパケットに指定された書込みアドレスを再変換して用いている。

【0022】本発明による第2のプロセッサ間通信方法は、複数のプロセッサ各々がネットワークによって結合される並列コンピュータのプロセッサ間通信方法であって、宛先プロセッサでのパケット受信時に論理アドレスで指定された前記宛先プロセッサでの書込みアドレスを物理アドレスに変換するステップと、その変換結果を記憶するステップと、記憶している変換結果が使えなくなる前に次に必要となるアドレス変換を先行して行うステ

ップとを前記複数のプロセッサ各々に備えている。

【0023】本発明による第3のプロセッサ間通信方法は、複数のプロセッサ各々がネットワークによって結合される並列コンピュータのプロセッサ間通信方法であって、宛先プロセッサへのデータ送信時に論理アドレスで指定された送信データの読出しアドレスを物理アドレスに変換するステップと、その変換結果を記憶するステップと、記憶している変換結果が使えなくなる前に次に必要となるアドレス変換を先行して行うステップとを前記複数のプロセッサ各々に備えている。

【0024】本発明による第4のプロセッサ間通信方法は、複数のプロセッサ各々がネットワークによって結合される並列コンピュータのプロセッサ間通信システムであって、宛先プロセッサでのパケット受信時に論理アドレスで指定された前記宛先プロセッサでの書込みアドレスを物理アドレスに変換するステップと、その変換結果を記憶するステップと、記憶している変換結果が使えなくなる前に次に必要となるアドレス変換を先行して行うステップと、宛先プロセッサへのデータ送信時に論理アドレスで指定された送信データの読出しアドレスを物理アドレスに変換するステップと、その変換結果を記憶するステップと、記憶している変換結果が使えなくなる前に次に必要となるアドレス変換を先行して行うステップとを前記複数のプロセッサ各々に備え、前記送信データを送るパケットにおいて前記書込みアドレスを指定しておき、先に送られたアドレス変換を依頼するパケットに付加された書込みアドレスと前記送信データを送るパケットに指定された書込みアドレスとが異なる時に前記送信データを送るパケットに指定された書込みアドレスを再変換して用いるようにしている。

【0025】また、本発明による第4のプロセッサ間通信方法では、1つの通信コマンドで複数のパケットによってデータが送られる場合、前記アドレス変換を依頼するパケットをそのコマンドの最初のデータを送るパケットの前だけに送り、その後に前記データを送るパケットだけを送るようにしている。

【0026】さらに、本発明による第4のプロセッサ間通信方法の変換結果を記憶するステップは、送り元プロセッサ毎に変換結果を管理し、同じ送り元プロセッサに対して現在使用中の変換結果と先行してアドレス変換した結果との2つの変換結果を記憶するようにしている。

【0027】さらにまた、本発明による第4のプロセッサ間通信方法のアドレス変換を要求するパケットにそのコマンドで送る総データ長の受信側でのページ数を指定して送り、前記宛先プロセッサにおいてそのページ数をカウントダウンすることで前記コマンドで書かれる最後のページを予測し、前記最後のページ以降の無駄な先行するアドレス変換を抑止するようにしている。

【0028】すなわち、本発明の第1のプロセッサ間通信システムでは、宛先プロセッサでの書込みアドレスが

10

20

30

40

50

論理アドレスで指定されている場合、宛先プロセッサでの書込みアドレスのアドレス変換を先行して行うためのアドレス変換を依頼するパケットを送っている。また、読出しアドレスも論理アドレスで指定されている場合には、このアドレス変換を依頼するパケットを、読出しアドレスを変換する前に送って先行するアドレス変換を早く始められるようにしている。

【0029】本発明の第2のプロセッサ間通信システムでは、続けて送られてくるデータの書込みアドレスのアドレス変換を先行してできるようにし、1つのコマンドのデータ転送中に発生するアドレス変換を事前に行うことによって、宛先プロセッサでの書込みアドレスのアドレス変換のオーバーヘッドを隠蔽している。

【0030】本発明の第3のプロセッサ間通信システムでは、読出しアドレスも論理アドレスで指定されている場合、読出しアドレスのアドレス変換を先行してできるようにすることで、1つのコマンドのデータ転送中に発生する読出しアドレスのアドレス変換のオーバーヘッドを隠蔽している。

【0031】本発明の第4のプロセッサ間通信では、読出しアドレス、書込みアドレスのアドレス変換を先行して行えるようにし、1つのコマンドのデータ転送中の読出しアドレス及び書込みアドレスのアドレス変換のオーバーヘッドを隠蔽している。また、1つのコマンドが複数のパケットで送られる場合には、書込みアドレスのアドレス変換を依頼するパケットを最初の1回だけ送り、書込みアドレスのアドレス変換をデータが到着する前に行えるようにしている。さらに、書込みアドレスのアドレス変換結果は送り元プロセッサ毎に管理し、あるプロセッサが送ったパケットの書込みアドレスのアドレス変換結果が、他のプロセッサからのパケットによって壊されないようにしている。さらにまた、アドレス変換を依頼するパケットに書込み領域のページ数を付加して送り、先行して行われる書込みアドレスの変換が最後のページで止まり、不必要な先行アドレス変換を行わないようにしている。これによって、必要なアドレス変換が無駄なアドレス変換に邪魔されることを防いでいる。

【0032】上述したように、送り側プロセッサの送信装置が送るデータの読出しアドレスのアドレス変換やデータ読出しを行う前に、宛先プロセッサに書込みアドレスのアドレス変換を依頼するパケットを送ることによって、先行して書込みアドレスのアドレス変換を行い、データ到着時にアドレス変換の必要がないので、書込みアドレスを論理アドレスで指定したプロセッサ間通信において、宛先プロセッサの受信装置でのアドレス変換の時間を削減することが可能となる。

【0033】また、本発明を用いない場合、宛先プロセッサの受信装置のアドレス変換の時間を短くするには、アドレス変換テーブルを受信装置内に全て保持しておく必要があるが、本発明を用いると、アドレス変換を先行

して行い、その結果だけを記憶すれば、データ到着時にはその結果を用いることが可能となるので、アドレス変換結果を保持するための記憶手段の容量を少なくすることが可能となる。

【0034】さらに、1つのコマンドで送られるデータが複数のページになる場合、読出しアドレスのアドレス変換、書込みアドレスのアドレス変換を先行して行うことによって、プロセッサ間の通信がアドレス変換によって中断されることなく行われる。したがって、受信装置でのアドレス変換のオーバーヘッドの削減や隠蔽を図ることが可能となる。

【0035】

【発明の実施の形態】次に、本発明の一実施例について図面を参照して説明する。図1は本発明の一実施例によるプロセッサ間通信システムの構成を示すブロック図である。図1において、本発明の一実施例によるプロセッサ間通信システムは複数のプロセッサ1-1~1-nがネットワーク100に接続された並列コンピュータである。プロセッサ1-1~1-nはそれぞれCPU2と、主記憶装置3と、送信装置4と、受信装置5とから構成されている。

【0036】送信装置4はコマンド読出し回路41と、コマンドレジスタ42と、パケット生成回路43と、アドレス変換回路44と、FIFO(First In First Out)メモリ45とから構成されている。

【0037】コマンド読出し回路41はCPU2からの送信要求によって、主記憶装置3上に設けられたコマンドキュー31からコマンド31aを読出し、コマンド31aの内容を検査し、エラーがなければ、コマンドレジスタ42に格納する。

【0038】パケット生成回路43はコマンドレジスタ42内のコマンド31aに書かれた読出しアドレスをアドレス変換回路44を使って論理アドレスから物理アドレスに変換する。また、パケット生成回路43はコマンドレジスタ42内の宛先プロセッサ等の情報を使ってパケットのヘッダを生成する。

【0039】さらに、パケット生成回路43はアドレス変換回路44によって変換された物理読出しアドレスからのデータの読出しを主記憶装置3に依頼し、読出した送信データ32をパケットとしてネットワーク100に送出す。さらにまた、パケット生成回路43はコマンドレジスタ42内の宛先プロセッサの書込みアドレスからアドレス変換パケットを生成し、ネットワーク100に送出す。

【0040】アドレス変換回路44はコマンドレジスタ42内の読出しアドレスを論理アドレスから物理アドレスに変換する。その際、アドレス変換回路44は主記憶装置3上のページテーブル33の情報を参照してアドレス変換を行う。ネットワーク100とパケット生成回路



43との間にはFIFOメモリ45があり、データ幅の変換やエラー検出符号等の作成を行う。

【0041】受信装置5はFIFOメモリ51と、送り元プロセッサ番号レジスタ52と、書込みアドレスレジスタ53と、アドレス変換回路54と、データ書込み回路57と、総ページ数カウンタ58とから構成されている。

【0042】FIFOメモリ51はネットワーク100から送られてきたパケットを一時的に格納する。送り元プロセッサ番号レジスタ52はパケットヘッダに含まれる、そのパケットを送信したプロセッサのプロセッサ番号を格納する。書込みアドレスレジスタ53はパケットヘッダに含まれている、パケットのデータの書込みアドレスを格納する。

【0043】アドレス変換回路54は書込みアドレスレジスタ53内の書込みアドレスを論理アドレスから物理アドレスに変換し、データ書込み回路57に物理書込みアドレスを渡す。アドレス変換回路54内には変換テーブル55があり、論理ページ番号55aと物理ページ番号55bとの対応表が格納されている。この対応表は主記憶装置3上のページテーブル33の値を参照して書換えられる。また、アドレス変換回路54は送り元プロセッサ番号レジスタ52内の送り元プロセッサ番号を使って変換テーブル55の論理アドレス55aを読み出し、その値と書込みアドレスレジスタ53内の論理ページ番号59とを比較するための比較器56を持つ。

【0044】データ書込み回路57はパケットで送られてきた受信データ34を、主記憶装置3上の、アドレス変換回路54から出力された書込みアドレスに書込む。総ページ数カウンタ58は1つのコマンドで送られるデータの書込み側でのページ数を格納し、1つのページをアドレス変換する毎に1ずつ減らしていき、最後のページ以降の先行するアドレス変換を行わないようにする。総ページ数カウンタ58は送り元プロセッサ毎にそれぞれのページ数をカウントする。

【0045】図2(a)は本発明の一実施例で用いられるアドレス変換パケットの構成を示す図であり、図2(b)は本発明の一実施例で用いられる通常のパケットの構成を示す図である。すなわち、図2には本発明の一実施例によるプロセッサ間通信システムで通信される2種類のパケットの形式を示している。

【0046】アドレス変換パケットAは宛先プロセッサ番号A1と、送り元プロセッサ番号A2と、書込み論理アドレスA3と、総転送ページ数A4とから構成されている。宛先プロセッサ番号A1はネットワーク100内のルーティングに使われ、総転送ページ数A4は1つのコマンドで送る書込み側での総ページ数である。

【0047】通常のパケットBは宛先プロセッサ番号B1と、送り元プロセッサ番号B2と、データ長B3と、書込み論理アドレスB4と、データB5とから構成され

ている。データ長B3はこのパケットで送るデータの長さである。

【0048】図3は図1の変換テーブル55の詳細な構成を示す図である。図3において、変換テーブル55は各送り元プロセッサ番号毎に2つの論理ページ番号552、555と物理ページ番号553、556との組みを格納することができる。それぞれにはその物理ページ番号が有効かどうかを示すビット554、557が付加されている。また、それぞれの送り元プロセッサ番号の組みのどちらが現在使われている情報かを示すビット551を持つ。先行してアドレス変換した結果はこのビット551で現在使われていない方に格納される。

【0049】変換テーブル55から送り元プロセッサ番号601によって選ばれた2つの論理ページ番号は2つの比較器56a、56bで、書込みアドレスレジスタ53からの論理ページ番号59と比較され、その結果がMUX(multiplexer)60で物理アドレスを選択するのに使われる。

【0050】通常のアドレス変換の処理は以下のようになる。送り元プロセッサ番号601によって選ばれた2つの論理ページ番号552、555を書込みアドレスレジスタ52からの論理ページ番号59と比較する。もし、同じものがあれば、物理アドレス553、556からMUX60によって選択してアドレス変換に使用される。

【0051】この時、物理ページ番号が有効でない場合にはその情報が主記憶装置3上のページテーブル33から読み出している途中なので、その読み出しを待って、アドレス変換に使用する。また、MUX60で選択したものがビット551で示された現在使用しているものと違う場合には、ビット551を書き替えておく。

【0052】もし、どちらの論理ページ番号552、555も異なる場合にはページテーブル33をアクセスする必要があるが、本発明を用いて先行してアドレス変換を行っている場合にはこのようなことは起こらない。

【0053】次に、先行して行うアドレス変換の処理について述べる。送り元プロセッサ番号601によって選ばれた2つの論理ページ番号552、555のうち、現在使っているのがどちらかを示すビット551で現在使っていない方に、書込みアドレスレジスタ53からの論理ページ番号59を書込む。その後、対応するその物理ページが有効かどうかを示すビット554、557のいずれかを無効にしてから、主記憶装置3上のページテーブル33をアクセスして物理ページ番号を読み出し、変換テーブル55の物理ページ番号553、556のいずれかに書込み、有効ビット554、557のいずれかを有効にする。

【0054】図4は本発明の一実施例による送り元プロセッサでのコマンドの処理手順を示すフローチャートであり、図5は本発明の一実施例による宛先プロセッサで

のパケットの処理手順を示すフローチャートである。これら図 1～図 5 を参照して本発明の一実施例によるコマンド及びパケットの処理手順について説明する。

【0055】送り元プロセッサでコマンド処理が行われる場合、プロセッサ 1-1 からはアドレス変換パケット A で宛先プロセッサに書込み論理アドレス A 3 と総転送ページ数 A 4 とが送信される。その後、プロセッサ 1-1 は読出しアドレスのアドレス変換を開始し（図 4 ステップ S 1）、読出しアドレスのアドレス変換が終了するのを待つ（図 4 ステップ S 2）。

【0056】プロセッサ 1-1 はアドレス変換が終了すると、1 パケット分のデータの読出しを行い（図 4 ステップ S 3）、パケットを送信する（図 4 ステップ S 4）。今送ったパケットが最後のパケットでなければ（図 4 ステップ S 5）、プロセッサ 1-1 は読出しアドレスの先行アドレス変換が必要かどうかを判断し（図 4 ステップ S 6）、もし、読出しアドレスの先行アドレス変換が必要であれば、読出しアドレスの先行アドレス変換を開始する（図 4 ステップ S 6）。

【0057】読出しアドレスの先行アドレス変換が必要かどうかは読出したアドレスがページ境界に近付いているかどうかで判断し、先行アドレス変換の結果が必要になるまでに先行アドレス変換が終了するようなタイミングで開始する。この処理を最後のパケットを送り終るまで続ける。

【0058】続いて、宛先プロセッサでパケット処理が行われる場合、プロセッサ 1-1 はパケットを受信すると、それがアドレス変換パケット A かどうかを調べる

（図 5 ステップ S 11）。コマンドの最初のパケットはアドレス変換パケット A なので、パケットに指定された書込みアドレスの先行アドレス変換を行う（図 5 ステップ S 12）。

【0059】そのうちに、通常のパケットが送られてくると（図 5 ステップ S 11）、プロセッサ 1-1 は先行アドレス変換が終了しているかどうかを調べ、終了するまで待つ（図 5 ステップ S 13）。アドレス変換が終了したかどうかは、図 3 に示す変換テーブル 5 の送り元プロセッサ番号と論理ページ番号とが同じ変換テーブルのエントリの物理ページ番号で有効になっているかどうかで判断する。

【0060】アドレス変換が終了すると、プロセッサ 1-1 はデータを書込み（図 5 ステップ S 14）、先行のアドレス変換が必要かどうかを調べる（図 5 ステップ S 15）。先行のアドレス変換は読出しアドレスのアドレス変換と同じように、先行のアドレス変換結果を使用するパケットが到着する前にアドレス変換が終了しているタイミングで行う。また、総ページ数カウンタ 58 が現在のページが最後のページでないことを示している場合にのみ先行アドレス変換を行う（図 5 ステップ S 16）。

【0061】このように、送り側プロセッサの送信装置 4 が送るデータの読出しアドレスのアドレス変換やデータ読出しを行う前に、宛先プロセッサに書込みアドレスのアドレス変換を依頼するパケットを送ることによって、先行して書込みアドレスのアドレス変換を行い、データ到着時にアドレス変換の必要がなくなるので、書込みアドレスを論理アドレスで指定したプロセッサ間通信において、宛先プロセッサの受信装置 5 でのアドレス変換の時間を削減することができる。

10 【0062】また、本実施例を用いない場合、宛先プロセッサの受信装置のアドレス変換の時間を短くするにはアドレス変換テーブルを受信装置内に全て保持しておく必要があるが、本実施例を用いた場合にはアドレス変換を先行して行い、その結果だけを記憶すれば、データ到着時にはその結果を用いることができ、アドレス変換結果を保持するための記憶手段の容量が少なくてもいいことである。

【0063】さらに、1 つのコマンドで送られるデータが複数のページになる場合、読出しアドレスのアドレス変換、書込みアドレスのアドレス変換を先行して行うことによって、プロセッサ間通信をアドレス変換によって中断することなく行うことができる。したがって、受信装置 5 でのアドレス変換のオーバーヘッドの削減や隠蔽を図ることができる。

【0064】

【発明の効果】以上説明したように本発明によれば、複数のプロセッサ各々がネットワークによって結合される並列コンピュータのプロセッサ間通信システムにおいて、複数のプロセッサ各々が、宛先プロセッサへのデータ送信時に論理アドレスで指定された宛先プロセッサでの書込みアドレスのアドレス変換を依頼するためのパケットを宛先プロセッサに送信し、送り元プロセッサから送られてきた書込みアドレスを物理アドレスに変換し、その変換結果を、アドレス変換を依頼するためのパケットの後に送られてくるデータの書込みアドレスとすることによって、受信装置でのアドレス変換のオーバーヘッドの削減や隠蔽を図ることができるという効果がある。

【図面の簡単な説明】

40 【図 1】本発明の一実施例によるプロセッサ間通信システムの構成を示すブロック図である。

【図 2】（a）は本発明の一実施例で用いられるアドレス変換パケットの構成を示す図、（b）は本発明の一実施例で用いられる通常のパケットの構成を示す図である。

【図 3】図 1 の変換テーブルの詳細な構成を示す図である。

【図 4】本発明の一実施例による送り元プロセッサでのコマンドの処理手順を示すフローチャートである。

50 【図 5】本発明の一実施例による宛先プロセッサでのパケットの処理手順を示すフローチャートである。

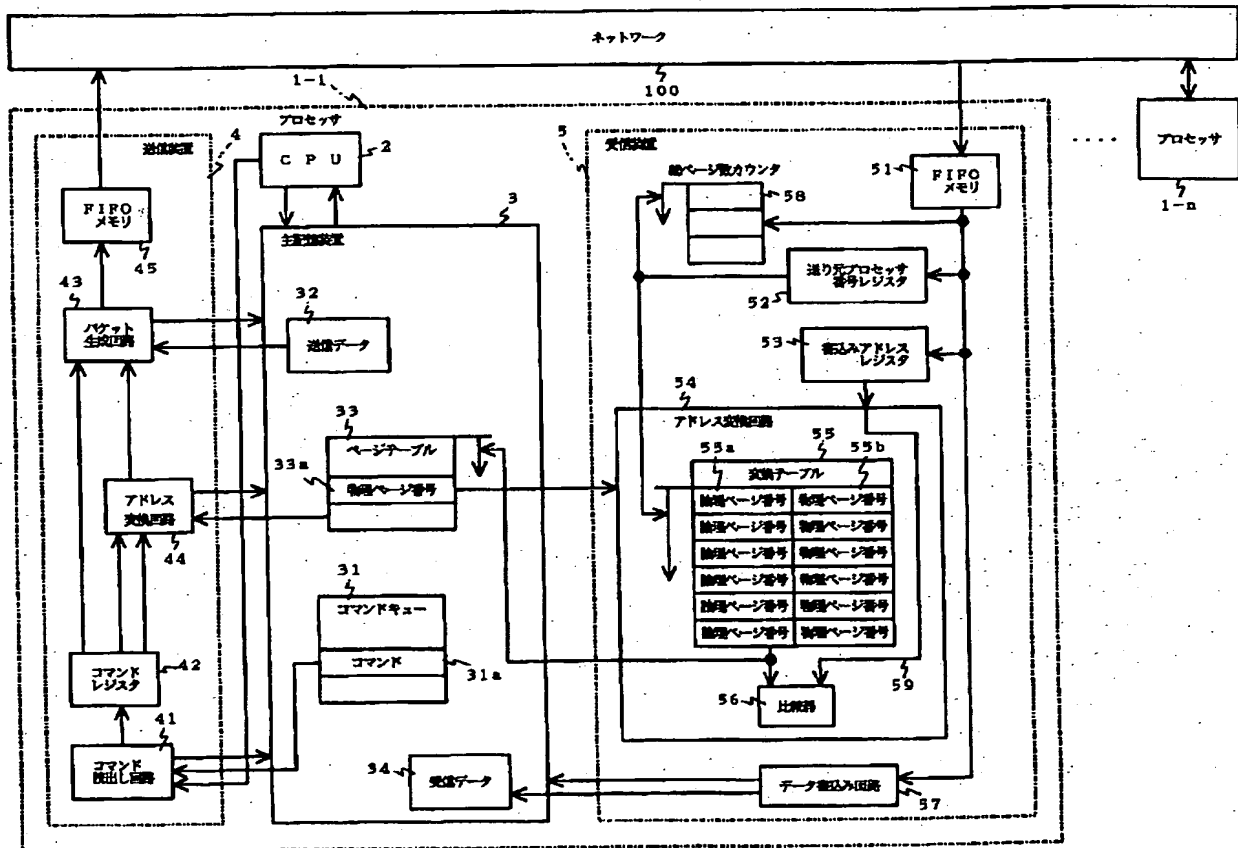
## 【符号の説明】

- 1-1～1-n プロセッサ  
 2 CPU  
 3 主記憶装置  
 4 送信装置  
 5 受信装置  
 31 コマンドキュー  
 31a コマンド  
 32 送信データ  
 33 ページテーブル  
 33a, 55b, 553, 556 物理ページ番号  
 34 受信データ  
 41 コマンド読出し回路  
 42 コマンドレジスタ  
 43 パケット生成回路  
 44, 54 アドレス変換回路  
 45, 51 FIFOメモリ  
 52 送り元プロセッサ番号レジスタ  
 53 書き込みアドレスレジスタ

## \* 55 変換テーブル

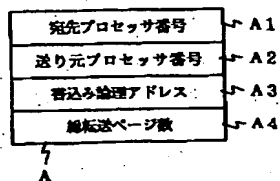
- 55a, 552, 555 論理ページ番号  
 56, 56a, 56b 比較器  
 57 データ書込み回路  
 58 総ページ数カウンタ  
 59 書き込みアドレスレジスタからの論理ページ番号  
 60 MUX  
 100 ネットワーク  
 501 現在使っているのが0か1かを示すビット  
 10 554, 557 物理ページ番号が有効かどうかを示すビット  
 601 送り元プロセッサ番号  
 A アドレス変換パケット  
 A1, B1 宛先プロセッサ番号  
 A2, B2 送り元プロセッサ番号  
 A3, B3 書き込み論理アドレス  
 A4 総転送ページ数  
 B 通常のパケット  
 \* B5 データ長

【図1】

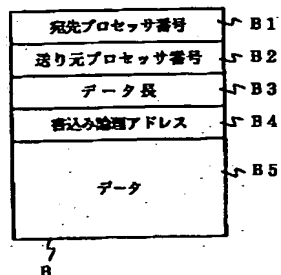


【図2】

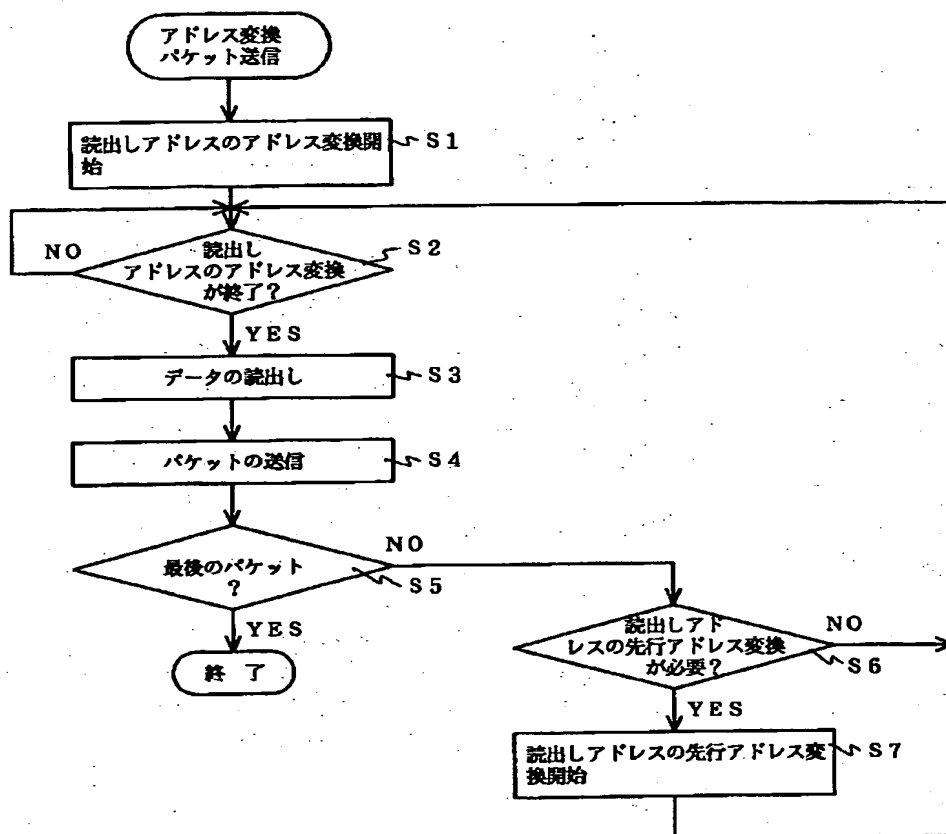
(a) アドレス変換パケット



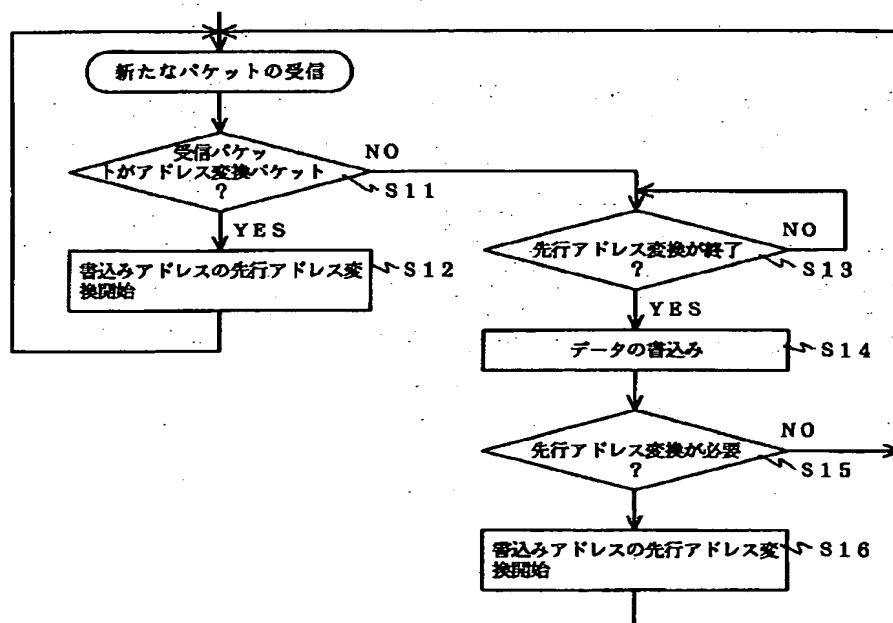
(b) 通常のパケット



【図4】



【図5】



【図3】

